

System DigiTool a budoucnost dlouhodobé ochrany DD

Andrea Fojtu
andrea.fojtu@ruk.cuni.cz

Setkání uživatelů DSpace, Ostrava, 15.4.2009

DigiTool (DTL)

- nástroj na správu digitálních sbírek (digital assets management)
- svojí funkcí mají DTL a Dspace mnoho společného
 - díky svojí open-source podstatě (resp. komunitě) jsou některé prvky propracovanější
 - porovnání:
 - <http://www.rsp.ac.uk/software/surveyresults>

DigiTool se mění na Rosetta

- Rosetta - výsledek spolupráce NLNZ a Ex Libris (i DDB a KB - zkušenosti s DIAS)
 - DTL + DPS (Preservation) = ROSETTA
 - open platform, nikoliv open-source
 - postaven na principu modelu OAIS
 - zaměřen především na dlouhodobou ochranu DD

Dlouhodobá ochrana DD

- digitální data
 - výsledek digitalizace
 - “digital-born”
- jsou uchovány v digitálním repozitáři, resp. archivu (např. DSpace)
- digital documents may last forever - or five years, whichever comes first (Rothenberg, 1997)

Deposit (SIP)

- různé producenti dat
 - jeden producent - 1 a více tzv. agentů
 - proces vkládání (depositu) v modulu Deposit končí registrací SIP do systému --> do tzv. pracovní prostor (work/staging area), kde se se SIPem dále pracuje
 - na deposit modul repozitáře lze napojit jakékoliv jiné aplikace
 - Web Curator Tool (NLNZ)

Pracovní prostor

- vložení --> validace --> schválení --> obohacení --> archivace
- identifikace formátu (validace JHOVE)
- kontrola integrity souborů (SHA a MD5) dodaných od producentů + antivirová kontrola (plug-in)
- obohacení metadaty, např. popisných
- migrace do požadovaného formátu – lze nejen pro samostatné soubory, ale i logické sbírky

AIP

- METS kontejnery pro všechny AIP
- datový model založen na formátu PREMIS
- nativní podpora Handle a URN:NBN
- všechny data na disku!
 - NE pásky – důvodem je „exit policy“
- „write once“ repository – je velmi velmi těžké něco smazat

DIP

- přístup - nástroj na doručení DD a nástroj na výhledávání DD (objektů)
- vyhledávání pomocí vlastního rozhraní, nebo tzv. “bundled Primo”

DO

- identifikace formátů --> nastavení rizik pro tyto formáty --> ohodnocení --> akce --> evaluace (podobné projektu Planets)
- risk management modul ukáže, které formáty jsou v ohrožení a kolik jich je
- lze připojit emulátory přes API (např. z projektu KEEP)
- reporty - BIRT (<http://eclipse.org/birt>)

DO #2

- systém ve výsledku odpovídá pravidlům důvěryhodného uložště (podle kritérií TRAC a DRAMBORA)
- technologie nejsou spasitelská řešení!!
 - klíčovou roli sehrávají lidské a finanční zdroje a management rizik
 - důvěryhodný repozitář ve výsledku představuje nikdy nekončící proces správy, testování a auditu (resp. certifikace)

Děkuji za pozornost!

andrea.fojtu@ruk.cuni.cz